

EXAMINING THE CLINICAL UTILITY OF RESEARCH DOMAIN CRITERIA
IN AN OUTPATIENT SAMPLE

Patrick K. Love

Thesis Prepared for the Degree of
MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

May 2018

APPROVED:

Jennifer Callahan, Major Professor
Randall J. Cox, Committee Member
C. Edward Watkins, Committee Member
Vicki Campbell, Chair of the Department of
Psychology
David Holderman, Dean of the College of
Liberal Arts and Social Sciences
Victor Prybutok, Dean of the Toulouse Graduate
School

Love, Patrick K. *Examining the Clinical Utility of Research Domain Criteria in an Outpatient Sample*. Master of Science (Psychology), May 2018, 47 pp., 3 tables, 2 figures, references, 68 titles.

This study examined the clinical utility of the recently released National Institute of Mental Health's (NIMH) research domain criteria (RDoC) by replicating and extending earlier work by using a demographically novel sample. Information retrieval and natural language processing of archival clinical records was used to achieve two main objectives: (1) estimate how well the RDoC domains match language used by clinicians by creating domain scores and (2) examine the differences between the *DSM*'s and RDoC's ability to predict treatment outcome using these domain scores and *DSM* diagnoses. The social systems RDoC category was found to be the strongest predictor of treatment outcome across all diagnostic measures.

Copyright 2018

by

Patrick K. Love

TABLE OF CONTENTS

	Page
LIST OF TABLES AND FIGURES.....	iv
CHAPTER 1. OVERVIEW OF MENTAL ILLNESS CLASSIFICATION SYSTEMS.....	1
CHAPTER 2. DETAILED REVIEW OF CLASSIFICATION SYSTEMS OF MENTAL ILLNESS.....	7
Criticisms of the <i>DSM</i>	9
An Overview of RDoC	10
RDoC Domains and Constructs	12
Negative Valence Systems.....	12
Positive Valence Systems	13
Cognitive System	15
Systems for Social Processes	16
Arousal/Regulatory Systems.....	17
Clinical Utility of RDoC.....	18
CHAPTER 3. METHOD	21
Participants.....	21
Measures	22
Outcome Questionnaire 45.2 (OQ).....	22
Psychiatric Diagnostic Screening Questionnaire (PDSQ)	23
Procedures	23
CHAPTER 4. RESULTS	26
CHAPTER 5. DISCUSSION.....	31
REFERENCES	42

LIST OF TABLES AND FIGURES

Page

Tables

Table 1. Demographics	37
Table 2. Number of Clients Who Met or Surpassed the PDSQ Cutoff Criteria by Diagnostic Category (N = 337)	38
Table 3. Fit Statistics of eRDoC and PDSQ GMM Models	39

Figures

Figure 1. GMM model with eRDoC scores predicting OQ total scores over time.....	40
Figure 2. GMM model with PDSQ scores predicting OQ total scores over time.	41

CHAPTER 1

OVERVIEW OF MENTAL ILLNESS CLASSIFICATION SYSTEMS

Using an accurate diagnostic system directly affects both the psychotherapy process and outcome by allowing clinicians to provide the most efficacious treatment for their client's disorder (Blashfield & Burgess, 2007; Insel, 2014). For example, a client misdiagnosed as having a depressive disorder when they actually have bipolar disorder may receive harmful and ineffective treatment like being prescribed anti-depressants or stimulants. Furthermore, accurate nosologies allow for improved communication among clinicians thus improving between-clinician treatment fidelity and treatment outcomes (Insel, 2014).

Unfortunately, the current de facto classification system, the *Diagnostic and Statistics Manual (DSM; APA, 2014)* suffers from high rates of co-morbid diagnoses (Clarkin & Kendall, 1992; Kessler, et al., 2012) and low between clinician diagnostic reliability (Regier, et al., 2013). Taken together, these findings suggest it may be an inadequate nosology. As a result, the National Institute of Mental Health (NIMH) has proposed research domain criteria (RDoC) that integrate accumulated evidence regarding the etiology of mental illness into a sophisticated, new multidimensional model of human behavior that may overcome some of the shortcomings associated with the DSM (Insel et al., 2010; NIMH, 2013). However, as noted by NIMH, the clinical utility of RDoC is largely unexplored and research in this area is still needed. Additionally, it is unclear how drastically clinicians would have to change the way they discuss mental illness should RDoC be adopted.

The goals of the study were to (1) examine the extent to which the competing DSM and RDoC frameworks map onto the natural language used by clinicians in their intake, progress, and

discharge summaries and (2) determine if the RDoC framework offers an improvement over the *DSM* framework with respect to predicative validity, as measured by psychotherapy outcomes.

Improvements in psychotherapy treatment outcomes are strongly associated with improvements made in the psychiatric classification system. Prior to the release of the first edition of the *DSM* (APA, 1952), a smaller percentage of clients treated with psychotherapy had positive outcomes than clients who received no psychotherapy treatment (Eysenck, 1952). The release of the theoretically neutral *DSM-III* (APA, 1980) allowed researchers to directly compare the efficacy of different treatment modalities in randomized clinical trials (Lilienfeld, Smith, & Watts, 2013) and a plethora of psychotherapy outcome studies have since emerged, ultimately serving as a catalyst for the evidence-based treatment movement and improved psychotherapeutic outcomes (Cuijpers, van Straten, Anderson, & Oppen, 2008). Arguably, the penultimate treatment of psychological distress may be personalized interventions. However, the *DSM*'s use of symptoms as the unit of analysis does not provided the level of specificity required for this to become a reality (Insel, 2104; Kapur, Phillips, & Insel, 2012).

RDoC was designed to be a classification framework that may provide a more precise etiological framework of psychological phenomena, eventually fostering personalized, evidence-based, treatments (Insel, 2014). The NIMH RDoC framework consists of five domains representing functional constructs of behavior (negative valence systems, positive valence systems, cognitive systems, social processes, and arousal/regulatory systems). These constructs are the summation of research regarding genes, molecules, circuits, physiology, behavior, self-reports and paradigms that underlie all normal and abnormal psychological behavior. In the RDoC matrix neural circuits serve as the primary unit of analysis much as behaviors serve as the primary unit of analysis for the *DSM*. While there has not yet been enough research to

definitively determine the clinical utility of RDoC, emerging research has yielded promising results. For example, RDoC criteria have been shown to be more accurate than the *DSM* in predicting the development of mood or disruptive disorders in children (Wakschlag, et al., 2015) and in discriminating between types of thought disorders (Clementz et al., 2015).

McCoy et al. (2015) conducted a study to determine the compatibility between the RDoC domains and the language used by practicing clinicians in electronic medical records. Not only did the study find that clinicians already routinely use language associated with RDoC domains, they also found that the closer the notes aligned with the RDoC domains the better they predicted treatment outcome as measured by the length of time before hospital readmission (McCoy et al., 2015). Additionally, the natural language used by clinicians was found to align with RDoC criteria and demonstrate greater accuracy in predicting treatment outcome, as compared to *DSM-IV* criteria and ICD9 codes. This project extends McCoy et al.'s (2015) study in a demographically novel sample, with outpatient psychotherapy outcomes as the focus instead of hospital readmissions, which may be more salient to medication stabilization needs of the inpatient sample.

Archival data from adults who participated in outpatient psychotherapy within a psychology clinic in the south central United States were used for the current study. In contrast, McCoy et al., (2015) used a sample primarily composed of inpatients from the northeastern United States. All clients in the archival set had previously consented to the use of their data for any future IRB and PCEC approved research studies. Approval from both the UNT Institutional Review Board (IRB; approval number: 15-410) and the Psychology Clinic Executive Committee (PCEC) that oversees research in the target clinic was secured before data collect began.

Data collected for this study included standardized data from clients as well as narrative data provided by clinicians in intake summaries, progress notes, and termination reports. The Outcome Questionnaire 45.2 (OQ) is routinely administered before the intake interview and prior to each subsequent psychotherapy session and was used to measure psychotherapy treatment outcome (Lambert, Gregersen, & Burlingame, 2004; Vermeersch et al., 2004). The Psychiatric Diagnostic Screening Questionnaire (PDSQ), a self-report questionnaire administered at intake was used to assess the presence of *DSM-IV* Axis I clinical symptoms across 13 clinical disorders (Zimmerman & Mattia, 2001).

Clinical notes included in this study included: intake reports, discharge summaries, and progress notes. Physical records of the intake notes and discharge summaries were copied and all identifying client information was redacted before being scanned into digital text files for analysis. In addition, electronic stored progress notes were retrieved via direct download from the electronic medical record (EMR) system onto a secure, HIPPA compliant external hard drive. Identifying information was immediately purged from the records. As a result, only de-identified records were available for analysis, minimizing the risk of breaching clients' confidentiality. Natural language processing and information retrieval techniques were used to calculate how well RDoC and *DSM-IV* criteria are captured in these clinical notes.

Estimated RDoC (eRDoC) domain scores were calculated from the free-text narrative notes following the method developed by McCoy et al. (2015). RDoC domains were defined using the symptom terms included in the NIMH workgroup concept matrix for each domain. The Microsoft Bing internet search engine application program interface (API) was used to conduct automated internet searches for each term used within an RDoC domain. The Bing API was used to automate the internet searches. The 50 most relevant web pages found within each term's

search were used to construct the RDoC corpus. Therefore, each RDoC domain's corpus were comprised of the 50 most relevant pages from each of its terms. Apache Tiika, version 1.4 (Mattmann & Zitting, 2006), and the Boilerpipe algorithm (Kohlschütter, Frankhauser, & Nejdl, 2010) were used to extract the text from each search result. Once the text has been extracted, Punkt (Kiss & Strunk, 2006) and Penn Treebank (Marcus, Santorini, & Marcinkiewicz, 1993) Natural Language Toolkit, version 2.0.4 (NLTK; Bird, Klein, & Loper, 2009) were used to preprocess the domain-defining corpus text into homogeneous bodies of text. These homogenous bodies of text were then partitioned into mono-, bi-, and trigrams along the tokenized boundaries. All sequences of text (n-grams) that contain stop words (words that do not add significant meaning to the analysis of the text, e.g. filler words) as defined by NLTK were dropped. The remaining bag-of-n-grams model were converted into a vector space model (Salton, Wong, & Wang, 1975) and the dimensionality of that model was reduced (Bradford, 2008) using latent semantic indexing (LSI; Deerwester, 1988) function of Gensim, version 0.8.9 (Rehurek, 2011). Since each document vector arose from a key concept in the workgroup matrix, the resulting model is inherently rooted in the five RDoC domains.

A mixed-effects, stepwise regression was used to determine eRDoC domain scores' and PDSQ cutoff scores associations with the treatment outcome, as measured by the change in OQ score between the first and last therapy sessions. Starting OQ score, and diagnoses was also controlled for in the model via entry into step one of the regression equation (via entry into step one of the regression equation). Additionally, two multinomial logistic regressions were conducted to determine if eRDoC domain scores or PDSQ category cutoff scores predict treatment outcome category as measured by the OQ. Finally, two growth mixture models were conducted, comparing the predictive validity of eRDoC loadings and DSM criteria as determined

by the Psychiatric Diagnostic Screening Questionnaire (PDSQ), for each session's OQ score longitudinally.

Based on the results from McCoy et al. (2015), it was hypothesized that eRDoC domain scores would have greater predictive validity of treatment outcome, as measured by OQ difference scores, than *DSM* diagnostic criteria as measured by PDSQ cutoff scores. It was also hypothesized that eRDoC scores of clinician intake reports would provide a better fitting model of treatment course trajectory than PDSQ cutoff scores.

CHAPTER 2

DETAILED REVIEW OF CLASSIFICATION SYSTEMS OF MENTAL ILLNESS

Despite multiple iteration and significant improvements, the *Diagnostic and Statistical Manual (DSM, APA, 2013)*, including its current iteration, still has considerable shortcomings. These include high rates of diagnostic comorbidity (Clarkin & Kendall, 1992; Kessler, et al., 2012; Maj, 2005), increased focus on splitting diagnoses into smaller categories (rather than exploring their etiology, for example), difficulty incorporating biological findings, and a uni-dimensional diagnostic model (Cloninger, 2009; Kendell & Jablensky, 2003; Marcus, Norris, & Coccaro, 2012; Meehl & Golden, 1982). This manuscript explores the short comings of the current *DSM* taxonomy followed by an exploration of the National Institutes of Mental Health's (NIMH) alternative multidimensional model of human behavior; the research domain criteria (RDoC).

An accurate psycho-diagnostic system is critical to advancing understanding of mental illness; perhaps as critical as examining etiological explanations of mental disorders (Lilienfeld, Smith, & Watts, 2013). Effective diagnostic systems should succinctly communicate relevant information about a client's condition, symptomology, and prognosis to other clinicians, researchers, managed care organizations, family members, the client themselves and others (Blashfield & Burgess, 2007). Furthermore, the overarching structure of the nosology helps contextualize a client's presenting problems within related and unrelated diagnostic categories (Lilienfeld, Smith, & Watts, 2013). Finally, a valid diagnostic system should provide professionals, both clinicians and researchers, additional information about the client's natural history, clinical profile, laboratory findings, family history and prognosis (Robins & Gauze, 1970). As just one example, reviewing the diagnostic history of major depressive disorder

(MDD) may serve as a useful illustration of how diagnostic criteria can shape both treatment and research of a given disorder.

The symptomology of MDD (feelings of extreme sadness and anhedonia that cause significant biological, cognitive, behavioral, and social disturbances) has remained remarkably stable since the syndrome was first described by the Greek physician Hippocrates in the 5th century BCE (Arnarson, & Craighead, 2013; Ritschel, Gillespie). However, the etiology and recommended treatments of MDD have varied markedly over time. Hippocrates postulated that melancholia (now known as MDD) was caused by an excess of black bile, one of the four humors present in the human body. The German psychiatrist Emil Kraepelin (1899) was the first to use the word depression to describe the behavior of clients being treated in German mental hospitals, and viewed the disorder as one end of the manic-depressive spectrum of mood disorders. Adolf Myer attributed the development of the disorder to a “depressive reaction” to negative life events. As such, Freud and other psychoanalytic clinicians argued the only cure for depression was to explore the psychosocial antecedent of the client’s depressed mood until the client achieved curative insight. Subsequently, behavioral (Skinner, 1953) and cognitive models (Beck, 1967; Beck 1976; Seligman, 1975) were proposed and with them came new methods of treatment (e.g., Beck’s cognitive behavioral therapy (CBT) focused on altering the underlying cognitions that maintain one’s depressed state). With so many competing theoretical orientations, the same client could be conceptualized very differently with resultant high variability in treatment approaches among possible providers.

The publication of the first edition of the *Diagnostic and Statistical Manual (DSM; 1952)* marked a significant change across the fields of psychiatry and psychology. However, the DSM and subsequent *DSM-II* were not without their own problems. Both editions contained short and

often vague descriptions of mental disorders with no clear explanation of how to diagnose them. Additionally, both manuals were heavily influenced by psychoanalytic approaches and written in such a way that it was difficult for other theoretical frameworks to follow the descriptions within. It was not until the third edition of the *DSM* (*DSM-III*) was published in 1980 that a more unified and trans-diagnostic methodology for diagnosing mental disorders emerged. Furthermore, the *DSM-III* was the first of the diagnostic manuals to use empirically supported and explicitly stated diagnostic criteria, a feature that persisted through *DSM-III-R*, *DSM-IV*, *DSM-IV TR* and *DSM-5*. Having a more theoretically neutral and unified description of mental disorders provided researchers and clinicians alike with a unified, theoretically neutral language (Wakefield, 1998) to communicate with.

Criticisms of the *DSM*

Despite iterative improvements, there are still many critics of the *DSM*. One criticism is that comorbidity rates present in the literature are much high than the rates we would expect to see from just chance alone. In fact, Boyd et al. (1984) suggested that the comorbidity rate of psychological diagnoses is as much as 22.4 times higher than what should be expected by chance. The fact that psychological comorbidity occurs more often than not suggests that comorbidity is an artifact of the current categorical diagnostic system (Maj, 2005).

A criticism levied at the most recent addition of the *DSM* is that it has over medicalized normality (Sommers & Satel, 2005). More specifically, the number of diagnoses has increased while the threshold criteria were also lowered for a number of existing diagnoses, potentially running a risk of reconceptualizing normative behaviors as pathological presentations in need of remedy (Bastra & Francis, 2012). Additionally, some authors have argued that *DSM-5*'s efforts

to achieve higher reliability has created a number of diagnostic categories which do not sufficiently capture the true breadth of the mental phenomena, a problem called the attenuation paradox (Clark & Watson, 1995; Loevinger, 1957; Skeem, Polascheck, Patrick, & Lilienfeld, 2011). Finally, there is growing evidence that the phenomenology of distress exists along a continuum rather than categorically (Cloninger, 2009; Kendell & Jablensky, 2003; Marcus, Norris, & Coccaro, 2012; Meehl & Golden, 1982). However, the *DSM-5* continues to ascribe a categorical nosological structure despite the increased external validity that is achieved when measuring mental disorders dimensionally (Craighead, Sheets, Craighead, & Madsen, 2011; Markon, Chmielewski, & Miller, 2011).

An Overview of RDoC

In response to many of the criticisms of the *DSM*, the National Institutes of Mental Health (NIMH) began working on an alternative multidimensional model of human behavior in 2010. The research domain criteria (RDoC) matrix (NIMH, 2013) was developed by an internal NIMH working group, in conjunction with a small number of expert external advisors. After an initial draft and public comment period, five individual constructs of cognition and behavior, called domains, were identified. In contrast to the *DSM*, RDoC takes a bottom up approach to classifying mental illness, using basic biological functions and neurological circuits to understand behavior (Kaufman, Gelernter, Hudziak, Tyrka, & Coplan, 2015). For a construct to be considered valid in RDoC, a body of data must exist in supporting that construct as a functional dimension of behavior or cognition (NIMH, 2013). Additionally, research has to have identified a specific neural circuit that accounts for the majority of variance implementing that particular function (NIMH, 2013). Contained within each construct is a definition of its

functional aspects, a summary of relevant neural circuitry, and its relationship with other constructs.

While there are currently five domains, each containing particular sub-constructs, the NIMH encourages researchers to explore the structure of other parts of cognition and behavior which have not yet obtained enough empirical support to be considered a valid domain. Additionally, the NIMH encourages comments and suggestions be made about the RDoC framework to address and resolve shortcomings in its structure and/or implementation. The intent is for the RDoC matrix to be a living document, changing more quickly to reflect advances in scientific knowledge than previous nosologies. However, RDoC was specifically created for research purposes and has yet to be validated as a way to formulate clinical diagnoses (NIMH, 2103). In spite of this, recent research has shown RDoC domains to be more predictive of treatment outcomes (McCoy et al., 2015), accurate in predicting eating disorder symptomology (Wildes and Marcus, 2015), and better at discriminating subtypes of psychosis (Cuthbert, 2014). Additionally, a recent study conducted by Wakschlag et al. (2015) examined temper loss in children using RDoC as a diagnostic framework and found it to more accurately predict the development of mood or disruptive disorders than similar *DSM* constructs. More importantly, risk of developing a disorder was significantly elevated, as much as 67%, at levels of temper loss that are considered normative by current criteria (Wakschlag et al., 2015). These findings highlight the advantages of using a truly dimensional framework and indicate that RDoC in its current form could be a more accurate psychopathological model than the traditional categorical system.

The rows of the RDoC matrix represent dimensions of behavior and related constructs are grouped into higher level groups called domains. The columns in the matrix represent units used

to measure each construct, beginning small and specific but gradually becoming more and more general. The current units represented in the columns are genes, molecules, cells, circuits, physiology, behavior, self-reports, and paradigms. While neurodevelopmental trajectories and interactions with the environment are not directly represented in the RDoC framework, they are still considered to be important areas of psychological research and can be parsed using the RDoC framework (NIMH, 2013). However, since development and the environment influence constructs differently over time and across units, these two were not formally included in the structure of the matrix to allow researchers more freedom while exploring these topics. It is also hoped that the influence of development and the environment on the RDoC domains and constructs will be considered in a systematic rather than granular fashion (NIMH, 2013).

RDoC Domains and Constructs

As previously mentioned, there are currently five domains, or higher-level constructs in the RDoC framework. These domains reflect the contemporary knowledge about the major systems of emotion, cognition, motivation, and social behavior. Each domain contains related sub-constructs which represent more granular aspects of that particular domain. While the constructs are grouped as follows, some artificial divisions had to be made. For example, neurological circuits and structures are often associated with multiple behaviors.

Negative Valence Systems

Negative valiance systems make up the first RDoC domain and are defined as primarily responsible for response to adverse stimuli (e.g. fear, anxiety, loss, etc.). The negative valiance systems domain contains the following constructs.

- Responses to acute threat (fear): Activation of the brain's defensive motivational system trigger protective behaviors. Normal fear is considered an adaptive response to conditioned or unconditioned stimuli. It can be in response to cognitive processing or internal representations as well as external stimuli and is affected by a variety of factors.
- Responses to potential harm (anxiety): This construct is the response to potential harm. More specifically, brain system are activated by potential, future harm. In contrast to fear, the threat is typically vague or ambiguous. Anxiety responses are characterized by increased vigilance, increased heart rate. Although similar, anxiety responses are qualitatively different than those of fear.
- Response to sustained threat: Responses to sustained threats reflect an adverse emotional state caused by chronic exposure to external stimuli from which escape would be adaptive. Exposure to the stimuli can be actual or anticipated. Additionally, the changes in cognition, behavior, and affect persist in the absence of the threat.
- Frustrative non-reward: Reactions in response to withdrawal/prevention of reward, typically occurring after sustained attempts to obtain it.
- Loss: A state occurring after being deprived of an emotionally significant stimulus. The stimuli may be social or non-social, concrete or non-concrete and may be episodic or sustained in nature.

Positive Valence Systems

The second RDoC domain, positive valence systems are primarily responsible for responses to positive stimuli such as reward seeking, consummatory behavior, and reward/habit learning. This domain contains the following constructs.

- Approach motivation: This is a multi-faceted construct that involves mechanisms or constructs influenced by pre-existing tendencies, learning, memory, stimulus characteristics, and deprivation states. Approach motivation may involve the assignment of incentive salience to stimuli. It contains the following sub-components:

- Reward valuation or the process computing the probability and benefits of perspective outcome in reference to available information
- Effort valuation/motivation to work or the process of computing the cost of obtaining the reward and the tendency to overcome response costs to obtain a reinforce
- Expectancy/reward prediction error which is triggered by exposure to internal or external stimuli, experiences, or contexts that predict the possibility of reward and can be altered by the experience of outcome and influence use of resources
- Action selection/preference-based decision making or the processes evaluating the cost/benefit analysis and decision making which occurs in the context of multiple potential choices

- Initial responsiveness to reward attainment: The processes associate with pleasant responses as displayed in subjective experiences, behaviors, and/or reward processing neural network engagement. This is the culmination of reward seeking behavior.

- Sustained/longer-term responsiveness toward attainment: These are the mechanisms associated with the termination of reward seeking behavior, e.g. satisfaction or satiation.

- Reward learning: A process which organisms acquire information about stimuli, action, and contexts that predict positive outcomes. This process explains how behavior is modified when novel rewards exceed expectations and is a type of reinforcement learning. A similar process may be involved in negative reinforcement learning.

- Habit: The repetitive physical or cognitive behaviors triggered by external stimuli that can go to completion without constant conscious oversight once initiated. The formation of

habits is frequently the consequence of reward learning. Expression of habits often become resistant to change.

Cognitive System

The third domain, these systems are responsible for various cognitive processes.

Cognitive systems contain the following constructs.

- Attention: Attention incorporates a wide range of processes that regulate access to higher level systems such as awareness, higher perceptual processes, and motor action. It has both limited capacity and limited completion ability (i.e., selective and divided attention).
- Perception: This is the process that interprets sensory data to construct and manipulate representations of the external environment, acquire information and make predictions about the external world to guide actions.
- Declarative memory: The acquisition, encoding, storage, and retrieval of cognitive representations of facts and events. Declarative memory provides substrate for episodic and semantic memory and facilitates the inferential and flexible extraction of new information from these relationships.
- Language: Language is a shared system of symbolic representations of the world, self, and abstract concepts that supports thought and communication between individuals.
- Cognitive control: This construct modulates the operation of other cognitive and emotional systems in the service of goal directed behavior when more immediate forms of response are not adequate for the current context. Active in novel contexts, cognitive control helps to select the most appropriate response from competing alternatives.

- Working memory: The process in which a limited capacity of information relative to a goal or task is held and momentarily updated. It may involve flexible binding of representations.

Systems for Social Processes

The fourth RDoC domain, there are systems that mediate responses to interpersonal settings and stimuli of various types and includes the following constructs.

- Affiliation and attachment: Affiliation is the engagement in positive social interactions while attachment is the selective affiliation as a consequence of social bonding. Both are behavioral consequences of social motivation and are moderated by social information processing, motivation, and the ability to recall social information. They include both positive and negative outcomes of social interaction. Clinical manifestations of disruptions include social withdrawal, social indifference, anhedonia, and over-attachment.

- Social communication: A dynamic process used for the exchange of socially relevant information and includes both receptive and productive aspects. The communication may contain receptive aspects that may be explicit or implicit. These aspects include facial recognition, affective recognition, and characterization among others. Productive aspects include eye contact, expressive reciprocation, and gaze following. The construct by its nature is interactive and reciprocal, often appearing very early in development. The underlying neural networks of social communication support both reflexive and volitionally controlled behavior. Social communication utilizes information from several modalities including facial, vocal, gestural, postural, and olfactory processing. It is organized into the following sub-constructs: reception of

facial communication, production of facial communication, reception of non-facial communication, and production of non-facial communication.

- Perception and understanding of others: Perception and understanding of others is the process and/or representations involved in being aware of, recalling information, reasoning, and/or making judgements about other animate entities. It is organized into the following sub-constructs:

- animacy perception (the ability to appropriately perceive that another entity is an agent)
- action perception (the ability to perceive the purpose of a performed action)
- and understanding mental states (the ability to make judgments about the intentions, beliefs, desires, and emotions of other animate entities with the added ability of making predictions or interpretations of their behaviors)

Arousal/Regulatory Systems

The fifth domain, arousal/regulatory systems are responsible for activating neural systems as appropriate for various situations, and for providing appropriate homeostatic regulation of such systems as energy balance and sleep. This domain is comprised of the three following constructs.

- Arousal: The sensitivity of an organism to stimuli, both internal and external. Arousal facilitates interaction with the environment and can be modulated by physical characteristics and motivational significance of the stimuli. It varies along a continuum that can be quantified by behavioral states. While arousal is a distinct phenomenon from motivation, they can and do interact with one another. Arousal may be associated with increased or decreased motor activity and can be regulated by homeostatic drives.

- Circadian rhythms: Circadian drives are endogenous self-sustaining oscillations that organize the timing of biological systems. They are synchronized to environmental cues, anticipate the environment, allow effective responses to challenges and opportunities, modulate homeostasis in the brain and body, and are evident across levels of organization (molecules, cells, circuits, systems, organisms, and social systems).
- Sleep and wakefulness: These are endogenous, behavioral states that reflect recurrent and coordinated changes in the functional organization of the brain, optimizing physiology, behavior, and health. Homeostatic and circadian process regulate both wakefulness and sleep. Sleep is reversible, is characterized by postural recumbence, behavioral quiescence, and reduced responsiveness, and has a complex architecture and a predictable cycle of NREM and REM states that have their own distinct neural substrates and EEG profiles. Intensity and duration of sleep is affected by homeostatic regulation and by wakeful experiences. Sleep is evident at the cellular, circuit and system levels. It also has restorative and transformative effects that optimize neurobehavioral functions during wakefulness.

Clinical Utility of RDoC

NIMH developed RDoC primarily as a tool for researchers and explicitly stated that it is not intended to be used as a diagnostic tool by clinicians at this time (NIMH, 2013). Indeed, critics of RDoC state that it's inherently reductionist view of behavior ignores the contextual factors that influence cognition and, as such, leave practicing clinicians with too few tools to use when treating clients (Lieblich, Castle, & Everall, 2015). Additionally, reworking the nosology of mental health using scientific evidence as its basis is an honorable goal. However, we have no way of knowing if the units of measurements chosen by the RDoC working group are the correct

constructs to use (Lieblich et al., 2015). Finally, they argue RDoC's diagnostic agnosticism, the lack of clear distinctions between "disordered" and "normal" behavior, any attempts at achieving construct validity while identifying disorders will be in due to the lack of clear targets (Wakefield, 2015). These limitations of RDoC, the critics conclude, lead to a noble yet flawed nosology that will ultimately fail to yield superior treatments for mental disorders.

Despite these criticisms early research exploring the clinical utility of RDoC have found it performed better than *DSM* criteria in several clinically relevant areas. For example, a recent study examining temper loss in children, which used RDoC as a diagnostic framework, found it to more accurately predict the development mood or disruptive disorders than similar *DSM* constructs (Wakschlag et al., 2015). More importantly, risk of developing a disorder was significantly elevated, as much as 67%, at levels of temper loss that are considered normative by current criteria (Wakschlag et al., 2015). These findings highlight the advantages of using a truly dimensional framework, like RDoC, as opposed to the more traditional categorical approach. A second study identified three neurobiologically unique psychosis subtypes which do not follow traditional diagnostic boundaries (Clementz et al., 2015). Even though these subtypes have unique underlying structures, there was significant overlap in the behavioral symptoms displayed by each subtype indicating that behavior can have multiple biological causes. In this reality, using a biologically driven nosology such as RDoC has the potential to drastically increase not only the reliability but the validity of clinical diagnosis. While these studies prove promising, their impact is not likely to be felt by practicing clinicians for some time.

A more immediate issue of adopting a new nosology system facing clinical researchers is the compatibility between old and new diagnostic systems. If the two nosologies are too discrepant, clinicians and clinical researchers would have to learn an entirely new vocabulary to

discuss cases. In addition, decades of clinical research may become irrelevant to the field. McCoy et al. (2015) conducted a study to determine the compatibility between the RDoC domains and the language used by practicing clinicians. Their study drew data from intake and discharge summaries within inpatient psychiatric departments of a large Northeastern hospital network. Not only did the study find that clinicians already routinely use language associated with RDoC domains, they also found that the closer the notes aligned with the RDoC domains the better they predicted treatment outcome as measured by the length of time before hospital readmission (McCoy et al., 2015). Additionally, the language used by clinicians which aligned with RDoC criteria were more accurate predicting treatment outcome when compared to DSM-IV criteria and ICD9 codes. While these findings suggest that the RDoC framework has clinical utility in its current form, the results have limited generalizability as they were found in a population of clients with more severe illness.

This study proposed to replicate the McCoy et al. study in a clinically and geographically distinct sample. More specifically, this study used archival records from an outpatient psychology training clinic in the south central United States. In addition to the sample population's cultural differences, this study's sample was known to differ from McCoy et al.'s in that the clinic primarily treats clients with depression and anxiety, as opposed to schizophrenia and bipolar disorder. The goals of this study were to (1) examine the extent to which these two frameworks map onto the language used by clinicians in their intake, progress, and discharge summaries and (2) determine if RDoC, offers an improvement in treatment outcome predictive validity over the *DSM*. To accomplish these aims, the methodology used in McCoy et al. (2015) was replicated in a clinically and demographically unique setting; a psychology training clinic located in the south-central region of the United States.

CHAPTER 3

METHOD

Participants

Archival data from adults who sought psychological treatment at the University of North Texas (UNT) Psychology Clinic were used for the current study. All services were provided by supervised trainee clinicians from one of the three accredited doctoral programs in the UNT Department of Psychology (i.e. clinical psychology, counseling psychology, or clinical health psychology, which is accredited as a clinical program). Student clinicians used a variety of treatment approaches with the clients including but not limited to cognitive-behavioral therapies, psychodynamic therapy, interpersonal therapy, acceptance and commitment therapy, and integrative therapies. Although some clients were treated with manualized therapy such as cognitive processing therapy, the majority of clients included in this study were not treated with manualized therapies.

During their intake session at the UNT Psychology Clinic, all clients are explicitly informed of the research mission of the clinic. More specifically, clients were informed that their de-identified treatment records could be used for certain studies approved by the UNT Institutional Review Board (IRB). All clients included in this study gave their informed consent for their outcome and diagnostic data to be used in research studies. The archival data was collected following procedures approved by the UNT IRB and the Psychology Clinic Executive Committee (PCEC).

The initial sample consisted of 644 adult individual psychotherapy clients who had attended at least two therapy sessions, had intake and termination reports available in their records, had progress notes available in the UNT Clinic electronic medical record (EMR)

system, had completed the OQ-45 prior to each of their therapy sessions and had completed the Psychiatric Diagnostic Screening Questionnaire (PDSQ). Three-hundred six of these clients started treatment at the psychology clinic prior to the clinic requiring all clients to complete the PDSQ. As a result, they completed the PDSQ in the middle of their treatment course and thus were excluded from the analyses as their PDSQ scores are not a true representation of their symptoms at intake. Prior to conducting any statistical analyses, these data were inspected for missing values, errors, and outliers. No systematic values of missing data were discovered.

Thus, the final sample consisted of 337 adult psychotherapy clients treated at the UNT psychology clinic. Clients completed an average of 16 sessions of therapy (median = 11 sessions), maximum number of 111 sessions. However, 64.1% of clients in the sample stopped therapy before the 16th session and 22.3% terminated after the third therapy session.

Measures

Outcome Questionnaire 45.2 (OQ)

The OQ, a 45-item self-report measure, is designed to assess a client's symptom severity over the previous week (Lambert, Gregersen, & Burlingame, 2004). OQ items ask clients to respond with how much they experienced various symptoms by responding on a 5-point Likert-scale, anchored at 0 and 4, with responses ranging from *almost never* to *always*. Answers to the OQ are then summed to yield a total score with a possible range of 0 - 180. Total raw scores greater than or equal to 63 indicate clinically significant distress. A 14-point reduction of total OQ score from the time of intake indicates a reliable improvement in symptom severity while a 14-point increase in total scores indicates reliable deterioration and worsening of symptom distress. The OQ has been used extensively in the literature to track therapeutic outcomes across

various treatments settings including psychology training clinics (Vermeersch et al., 2004), demonstrating high concurrent validity, specificity to change, and sensitivity to change during treatment (Vermeersch, Lambert, and Burlingame, 2010).

Psychiatric Diagnostic Screening Questionnaire (PDSQ)

The PDSQ is self-report questionnaire comprised of 111 dichotomous (“yes” or “no”) items assessing the respondent’s clinical symptoms for DSM-IV Axis I symptoms (Zimmerman and Mattia, 2001). Some PDSQ questions ask respondents to record symptoms experienced over the past six months while others inquire about symptoms experienced during the previous two weeks. This measure yields 13 clinical disorder subscale scores and one total raw scale score. Those 13 clinical disorders are major depressive disorder (MDD), post-traumatic stress disorder (PTSD), bulimia/binge eating disorder, obsessive-compulsive disorder (OCD), panic disorder psychosis symptoms, agoraphobia, social phobia, alcohol abuse/dependence, drug abuse/dependence, generalized anxiety disorder (GAD), somatization disorder, and hypochondriasis. Each subscale of the PDSQ includes a clinical cut off score as well as multiple critical items. Subscales on the PDSQ demonstrate good to excellent levels of internal consistency, test-retest reliability, and convergent validity. The PDSQ has a mean negative predictive validity of the subscales is 97% using cutoff scores with 90% sensitivity (Zimmerman and Mattia, 2001).

Procedures

At intake, clients completed the Psychiatric Diagnostic Screening Questionnaire (Zimmerman and Mattia, 2001) and a standardized demographic form which collects the client’s

age, self-identified race/ethnicity, gender, marital status, religion, occupation, income, high school graduation, and veteran status. See Table 1 for the demographics of the sample. See Table 2 for the frequency of clients who meet PDSQ cutoff criteria and fell in the clinical range. It should be noted that individual clients can meet criteria for multiple diagnoses by meeting multiple cutoff scores on the PDSQ. Therefore, the number of clients who met multiple PDSQ criteria is also included in Table 2.

Following the intake interview, clinicians composed an intake summary that included information on the presenting problem, the client's medical, social, educational, and psychological histories, as well as the clinician's initial clinical impressions. During the course of treatment, clinicians completed progress notes after each session which included a summary of topics discussed during the session, the clinician's impression of the client's level of functioning and symptom severity, any specific therapeutic interventions employed, and treatment plans for the upcoming sessions. Throughout the full course of treatment, clients were asked to self-report clinical symptoms experienced over the previous week by completing the OQ prior to each session. Upon termination, clinicians completed a discharge summary detailing the reason for terminating therapy, a summary of the course of treatment, their final clinical impressions, and any post treatment recommendations they had for the client. All records were signed by a supervising psychologist.

Archival data derived from clinical records were collected using two methods. First, physical records of the intake notes and discharge summaries were copied, and all identifying client information was redacted before these records were scanned into digital files for analysis. Second, electronically stored progress notes were retrieved via direct download from the UNT Psychology Clinic's electronic medical record (EMR) system onto a secure, HIPPA compliant

external hard drive. Any identifying information was immediately purged from the records. As a result, only de-identified records were available for analysis, minimizing the risk of breaching clients' confidentiality.

CHAPTER 4

RESULTS

Outcome Questionnaire 45.2 (OQ) score changes between the first and last therapy session were used to measure treatment outcome. OQ change scores were placed into four categories based on the magnitude and direction of change. The four outcome categories are: no reliable change as indicated by a change score less than 14; reliably improved as indicated by an OQ score reduction by 14 points or more; recovered as indicated by the participant's starting OQ score was in the clinical range, (scores of 63 or greater), has a 14 point or greater reduction, and their final OQ score is in the non-clinical range (scores of 62 or lower); and deteriorated as indicated by a 14 point or greater increase in OQ score between the start and end of treatment. Of the 337 participants, 53.8% fell with the no reliable change category, 20% fell within the reliable improvement category, 18.2% fell within the recovered category, and 7.9% fell within the deteriorated category.

Estimated research domain criteria (eRDoC) domain scores were calculated from the free-text narrative notes following the method developed by McCoy et al. (2015). RDoC domains were defined using the symptom terms included in the NIMH workgroup concept matrix for each domain. The Microsoft Bing internet search engine application program interface (API) was used to conduct automated internet searches for each term used within an RDoC domain. The Bing API was used to automate the internet searches. The 50 most relevant web pages found within each term's search were used to construct the RDoC corpus. Therefore, each RDoC domain's corpus were comprised of the 50 most relevant pages from each of its terms. Apache Tiika, version 1.4 (Mattmann & Zitting, 2006), and the Boilerpipe algorithm (Kohlschütter, Frankhauser, & Nejd, 2010) were used to extract the text from each search result.

Once the text was extracted, Punkt (Kiss & Strunk, 2006) and Penn Treebank (Marcus, Santorini, & Marcinkiewicz, 1993) Natural Language Toolkit, version 2.0.4 (NLTK; Bird, Klein, & Loper, 2009) were used to preprocess the domain-defining corpus text into homogeneous bodies of text. Monograms are one-word subsections of text, while bigrams contain two words and trigrams contain three. These groups are separated by normal boundaries, such as spaces or punctuation marks, which can be identified as tokens/typical boundaries. These homogenous bodies of text were partitioned into mono-, bi-, and trigrams along the tokenized boundaries to use as the bases of the NLTK analysis. All sequences of text (n-grams) that contain stop words (words that do not add significant meaning to the analysis of the text, e.g. filler words) as defined by NLTK were dropped. The remaining bag-of-n-grams model were converted into a vector space model (Salton, Wong, & Wang, 1975) and the dimensionality of that model was reduced (Bradford, 2008) using a latent semantic indexing (LSI; Deerwester, 1988) function of Gensim, version 0.8.9 (Rehurek, 2011). Since each document vector arose from a key concept in the workgroup matrix, the resulting model is inherently rooted in the five RDoC domains.

To score each narrative note, the text was transformed into the domain-defined vector space model. The similarity of the resulting clinical document vector to all of the RDoC concept vectors in the model was scored via cosine similarity (Singhal, 2001). The metadata on the search terms of origin for each document vector in the domain defined model, and thus each RDoC domain, were maintained to reduce the resulting clinical-document-versus-each-concept-document similarity scores. The metadata was used to partition the clinical-document versus-each-concept-document similarity scores for each clinical document grouped by RDoC domain of origin. Finally, five eRDoC similarity scores, one for each RDoC domain, was created for

each clinical document by aggregating and then calculating the mean per-domain similarity scores.

A mixed-effects, stepwise regression was used to determine Psychiatric Diagnostic Screening Questionnaire (PDSQ) cutoff scores and eRDoC domain scores' association with the treatment outcome, as measured by the change in OQ score between the first and last therapy sessions. Starting OQ score and diagnoses were controlled for in the model via entry into step one of the regression equation. The normality of residuals, homoscedasticity and linear relationship of the variables assumptions were not violated in this data as determined by examining a histogram and scatter plot of the residuals. Collinearity diagnostics were within normal limits for all variables. The alcoholism cutoff score of the PDSQ was found to be a significant predictor of OQ change score between first and last session ($\beta = -.13$) with the model explaining approximately 9% of the variance; $F(1, 334) = 17.815$; $p < .001$; $R^2 = .091$. In other words, clients who met the PDSQ alcohol dependence/abuse cut off score showed significantly less change over the course of treatment than clients who did not abuse alcohol.

Two multinomial logistic regressions were conducted to determine if eRDoC domain scores or PDSQ category cutoff scores were predictive of treatment outcome category as measured by the OQ. The eRDoC domain scores nor (Nagelkerke $R^2 = .069$, Model $\chi^2(15) = 21.81$, $p = .113$). The model which used PDSQ cutoff scores as predictors was statistically significant (Nagelkerke $R^2 = .211$, Model $\chi^2(39) = 71.39$, $p = .001$). However, the Pearson goodness-of-fit statistic was quite large and statistically significant ($\chi^2 = 753.74$, $p = .050$) indicating that the model of this data does not fit the data well, thus the model should not be interpreted (Long & Frees, 2006).

Next, two growth mixture models with linear trajectories were estimated separately using the structural equation model (SEM) framework. The SEM framework of growth mixture model (GMM) was used due to its greater tolerance of variance between measurement points. The first GMM model used eRDoC scores and the second model used PDSQ cutoff scores as time invariant predictors of OQ score over time using Mplus version 7 (Muthen & Muthen, 1998-2014). Treatment course was measured over up to 16 treatment sessions (the sample mean) to reduce problems with model convergence which occurred when the number of sessions was unconstrained. Robust FIML estimation was used to handle missing data due to attrition in therapy. This method is most appropriate because it uses all available data points while simultaneously accounting for data with non-normal distributions with adjustments to standard errors and scaling chi-squared statistics (Little et al., 2014).

Model fit was evaluated using the confirmatory fit index ($CFI \geq .90$), Tucker-Lewis index ($TLI \geq .90$), standardized root mean square residual ($SRMR \leq .08$), root mean square error of approximation ($RMSEA \leq .08$), and the value of RMSEA is within 90% confidence intervals (Weston & Gore, 2006). As chi-squared fit measures are highly influenced by sample size, the chi-square value was divided by the degrees of freedom in the model to correct for this bias, with corrected a ratio falling below 2 indicating good fit and a ration below 3 indicating acceptable fit (Schermelel-Engel & Moosbrugger, 2003). The models' Akaike information criteria (AIC) and Bayesian information criteria (BIC) statistics were also reported to aid in model fit comparison. The model with smaller AIC and BIC values is indicative of better fit to the data (Schermelel-Engel & Moosbrugger, 2003).

The final fit for both the eRDoC and PDSQ GMM models were adequate. See Table 3 for the fit statistics of both models. Examination of the model fit indices to determine which model

fit the data more closely yielded mixed results. The chi-square/degrees of freedom ratio, CFI and TLI model fit indices indicate the eRDoC model better fits the data while the SRMR, AIC and BIC fit statistics indicate that the PDSQ model better fits the data. The two models had similar RMSEA fit statistics. Thus, both models fit the data equally well and a further examination of their loadings is warranted.

In the eRDoC GMM model, there was no statistically significant difference between the starting OQ score and intake eRDoC scores. However, the systems for social processes eRDoC score had a statistically significant negative slope factor loading, indicating that higher eRDoC scores in this domain predicted decreases in OQ scores over time. The text from a sample of intake reports which had high eRDoC loadings on the systems of social process domain were reviewed to better determine the reason for this relationship. This qualitative review indicates that clients had higher loadings on this domain as they had more problems in this area of functioning. Thus, no other eRDoC scores had a statistically significant relationship with OQ trajectories. See Figure 1 for the eRDoC GMM model and loadings. No PDSQ cutoff scores significantly predicted the trajectory of OQ scores over time. However, meeting the cutoff score for MDD, drug abuse, alcohol abuse, GAD, Psychosis, Somatoform disorder or Hypochondriasis was associated with significantly having significantly higher OQ scores at the initial treatment session. See Figure 2 for the PDSQ GMM model and loading.

CHAPTER 5

DISCUSSION

This research was conducted with several goals in mind. First, to determine the clinical utility of the National Institutes of Mental Health (NIMH) research domain criteria (RDoC) classification of behavior and cognition by replicating and extending McCoy et al.'s (2015) study in a demographically and geographically unique sample. Second, determine if RDoC criteria improve upon the poor predictive validity of *Diagnostic and Statistical Manual* 4th ed. (*DSM-IV TR*) criteria in an actual client sample. *DSM-IV TR* criteria were used as a part of this study for two reasons. First, at the time of data collection the *DSM-5* had just been released and had not yet been fully implemented by the clinic that data were collected from. Second, *DSM-5* criteria were shown to have lower diagnostic consistency than *DSM-IV TR* criteria in field trials (Reiger, et al., 2012). Therefore, the use of *DSM-IV TR* may yield a more rigorous comparison of RDoC's ability to predict treatment outcome. Based on an emerging body of literature it was hypothesized that eRDoC scores would be better predictors of treatment outcome as measured by change in Outcome Questionnaire 45.2 (OQ) score between the first and last session as well as the growth trajectory of the client's symptoms over the course of treatment. Only one Psychiatric Diagnostic Screening Questionnaire (PDSQ) or estimated (eRDoC) category was found to predict OQ score change between the first and last therapy session.

The multiple analyses that were run yielded decidedly mixed results. First, neither nosological system was able to predict categorical treatment outcome as determined by change in OQ scores over time. However, some researchers argue against using categorical predictors of outcome as true treatment outcome is a continuous variable (Maxwell & Delaney, 1993; MacCallum, Zhang, Preacher, & Rucker, 2002). Thus, the categories are arbitrary and may not

be truly reflective of the effectiveness of therapy. That is why OQ total score change between first and last session was also examined as a continuous variable.

However, it does appear that the presence of alcohol use makes it much less likely that psychotherapy alone will be effective in reducing the clients' levels of distress. Clients who reported problematic alcohol use showed significantly less improvement in therapy than client's who did not report problematic alcohol use. This is a useful finding in isolation as it suggests clinician's need to treat alcohol use disorders before treating other psychological problems. Furthermore, this effect was large enough to overcome a nosy sample further highlighting its importance. This finding is consistent with several treatment studies which recommend treating problematic alcohol use before addressing comorbid disorders (Helzer, & Pryzbeck, 1988; Lynskey, 1998; Resick, Monson, & Chard, 2017).

While eRDoC scores and PDSQ cutoff scores as predictors of client's treatment growth curves yielded models that fit the data equally well, the two models demonstrated different relationships in the sample. For example, the PDSQ growth mixture model (GMM) model found clients who met the clinical cut off score for major depressive disorder (MDD), psychosis, somatoform disorder, and hypochondriasis had significantly higher OQ score at baseline than other diagnostic groups. This finding is interesting and suggests that these disorders cause a greater amount of subjective distress, it is very likely that this particular finding is sample dependent. More interestingly, this model was unable to predict the treatment trajectory of clients. Taken together, these findings indicate that *DSM* categories do identify patterns of behavior that are distressing to people, just not how they will feel long term.

The *DSM-IV TR* criteria's failure to predict treatment trajectory may be due in part to the large number of co-morbid diagnosis in the sample. When examining the data, 84.9% of the

sample met the criteria for two or more *DSM-IV TR* diagnoses. Thus, the muddled diagnostic makes it hard both clinically and statistically to make predictions about treatment outcome. For example, it is not hard to imagine that a client presenting with comorbid generalized anxiety disorder (GAD) and MDD will need therapy teaching different life skills than clients who present with OCD and bulimia. This finding strongly supports the argument that a categorical nosological system does not accurately capture psychopathology (Maj, 2005; Bastra & Francis, 2012). More specifically, this extremely high rate of comorbidity in *DSM-IV TR* criteria and the lack of predicative validity of either nosological system supports an argument that developmental psychopathologists have made, the principal of equifinality (Franklin, et al. 2105). The principal of equifinality states that the same observed behavior can have multiple underlying causes. For example, someone can appear sad and lose their appetite because they are grieving the loss of a loved one, have a negative core belief that they are worthless and therefore are not deserving of meals, have a neurochemical imbalance, are currently high on heroine, or just found a cockroach in their favorite soup. Effectively remedying the negative emotion and decreased appetite requires different interventions for each of these cases. For example, it may be extreme to send the person who had a roach in their soup to psychotherapy for depression. Similarly, it may be overkill to prescribe anti-depressants to a client currently experiencing normal, uncomplicated grief. Thus, the clinicians in this sample may have been treating symptoms which were ancillary to the cause of their problems.

In contrast, the GMM model which used eRDoC loadings to predict OQ total score changes over time did not identify any significant difference in start points. This is also problematic from a clinical utility stand point. A core function of a diagnostic system is to identify factors which cause a disruption in an individual's ability to function (Blashfield &

Burgess, 2007). Taken in isolation, this finding would suggest that RDoC domains have limited clinical use and would better function as a research tool. However, the fact that the eRDoC GMM model fit the data as well as the model using DSM criteria indicates it is at least as good as the DSM in predicting treatment outcome. This supports the idea that RDoC criteria have naturalistic clinical utility while being grounded in biological and genetic models of data.

Additionally, high eRDoC scores on the systems of social functioning domain predicted reduction in OQ total score over time. This is consistent with treatment outcome literature which indicates that social support is a strong predictor of treatment outcomes for a wide range of disorders (Cobb, 1976; Dadds & McHough, 1992; Longabaugh, Beattie, Noel, Stout, & Malloy, 1993; Steketee, 1993; Berkman, 2000; Dobkin, Civita, Paraherakis, & Gill, 2002; and Burgoyne, 2005). The exact relationship between social support and treatment outcome is less clear from this data set. The relationship be due to the fact that clients with strong social support networks do better in therapy, or it could be that psychotherapy is more effective at treating problems in social functioning. The literature on social support is evenly split with some examining the positive power of supportive social networks, (Longabaugh et al., 1993; Steketee, 1993; and Burgoyne, 2005) while others have shown positive treatment outcomes by improving the client's functioning within their social network (Koons, et al., 2001; O'hara, Stuart, Gorman, & Wenzel, 2000). Previous research has found sixteen sessions to be typical for delivery of an effective dose of therapy (for a summary review of the dose-effect literature that draws this conclusion, see Hansen, Lambert, and Froman, 2002) and thus insufficient dosage of therapy is not a likely explanation for the findings.

In an effort to better understand the nature of this relationship in this data set, the content of a portion of intake reports with high and low eRDoC loadings were examined. As expected,

intake reports with low eRDoC loadings on the systems of social functioning domain did not mention social functioning very much in any context. Notes with high eRDoC loadings in this domain mentioned problems in social systems as well as supportive social systems with equal frequency. Therefore, the strongest conclusions one can draw from this particular data set is clinicians who focused on social systems in their intake were more successful at reducing symptom severity over the first 16 weeks of treatment. However, this finding could shed light on how psychotherapy works across a broad range of symptoms and should be explored in future research.

While one of this study's strengths is the use of a naturalistic sample and the weekly recording of OQ total scores to measure symptom distress, the uncontrolled nature of the sample may have contributed to the poor predictive power of the models and its failure to perfectly replicate the findings in McCoy et al. (2015). More specifically, the OQ was intended to measure short term symptoms severity over a one-week period (Lambert, Gregersen, & Burlingame, 2004). As such, the data is much noisier than the outcome measure used in the McCoy study, days until rehospitalization. It could be that eRDoC domain loadings and *DSM IV-TR* criteria are able to predict outcome, just not to the fine degree predicting changes in OQ would require.

Even so, the results of this study are impressive and important. The fact that eRDoC loadings had equal, and in the case of social functioning better predictive power than *DSM* criteria, indicates that RDoC criteria have naturalistic clinical utility. By extension, this suggests that biologically based and dimensional models of disordered behavior are at least as accurate as more traditional categorical systems. Thus, it is highly likely that they can be improved to identify the etiology of both normal and maladaptive behavior. Additionally, the predictive

power of the systems of social functioning eRDoC score is especially impressive and is indicative of the strength of social process in client's psychological health.

Table 1

Demographics

Variable		Overall Sample
<i>Age, M (SD)</i>		27.4 (11.33)
<i>Household Family Income, M (SD)</i>		\$28,197.68 (\$40,882.74)
<i>Gender, n (%)</i>	Male	111 (32.9%)
	Female	225 (66.8%)
	Transgender	1 (0.3%)
<i>Ethnicity, n (%)</i>	White, Non-Hispanic	217 (64.4%)
	African-American	22 (6.5%)
	Latino(a)/Hispanic	38 (11.3%)
	Asian/Pacific Islander	14 (4.2%)
	Middle Eastern/Southeast Asian	7 (2.1%)
	Native American/Alaskan Native	5 (1.5%)
	Multiracial	23 (6.8%)
	Other/Did not identify	11 (3.3%)
<i>Marital Status, n (%)</i>	Single	273 (81%)
	Married	35 (10.4%)
	Divorced	18 (5.3%)
	Other	11 (3.3%)
<i>High School Diploma, n (%)</i>	Yes	320 (95%)
	No	16 (4.7%)
	Missing	1 (0.3%)
<i>Veteran Status, n (%)</i>	Yes	14 (4.2%)
	No	332 (95.5%)
	Missing	1 (0.3%)

Table 2

Number of Clients Who Met or Surpassed the PDSQ Cutoff Criteria by Diagnostic Category (N = 337)

Disorder	Number of Clients Who Meet Cutoff Score	Percent of Sample
MDD	120	35.6%
PTSD	132	39.2%
Bulimia/Binge Eating	41	12.2%
OCD	120	35.6%
Panic Disorder	135	40.1%
Psychosis	68	20.2%
Agoraphobia	137	40.7%
Social Phobia	230	68.2%
Alcohol Abuse/Dependence	85	25.2%
Drug Abuse/Dependence	55	16.3%
GAD	203	60.2%
Somatization Disorder	107	31.8%
Hypochondriasis	80	23.7%
Met No Cutoff Value	18	5.3%
Met One Cutoff Value	33	9.8%
Met Multiple Cutoff Values	286	84.9%

Table 3

Fit Statistics of eRDoC and PDSQ GMM Models

Fit Statistic	eRDoC GMM Model	PDSQ GMM Model
χ^2 , (<i>df</i> , <i>p</i> -value)	377.59 (199, .000)	600.19 (313, .000)
χ^2/df ratio	1.90**	1.92**
RMSEA (95% CI)	.05 (.043 – .059)**	.05 (.046 – .058)**
SRMR	.08*	.04**
CFI	.96*	.94
TLI	.96*	.94
AIC	26347.22**	26176.82**
BIC	26474.11**	26356.50**

Note: * Indicates adequate fit according to rules of thumb (Schermelleh-Engel & Moosbrugger, 2003; Weston & Gore, 2006). ** Indicates good fit according to rules of thumb (Schermelleh-Engel & Moosbrugger, 2003; Weston & Gore, 2006).

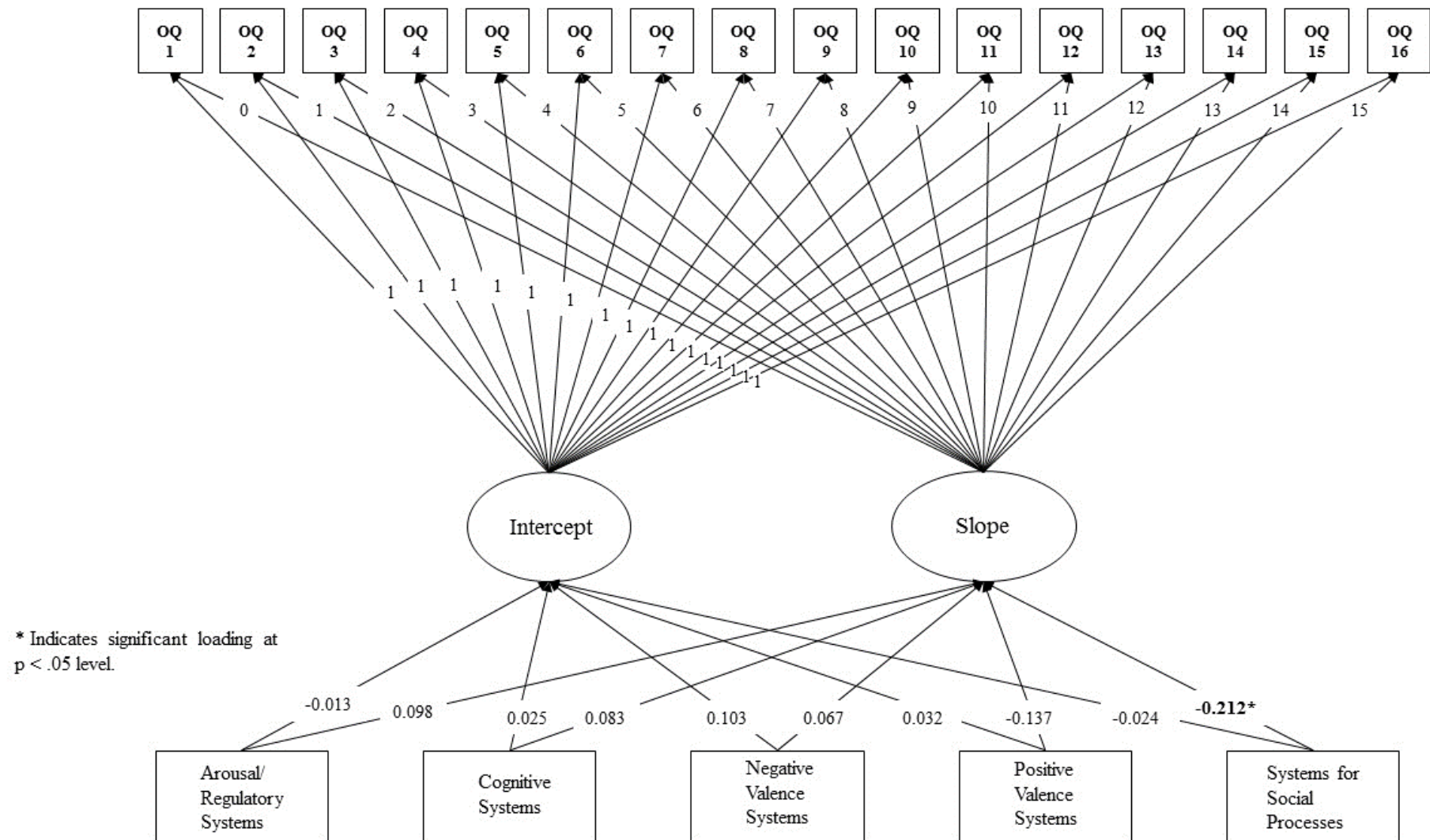


Figure 1. GMM model with eRDoC scores predicting OQ total scores over time.

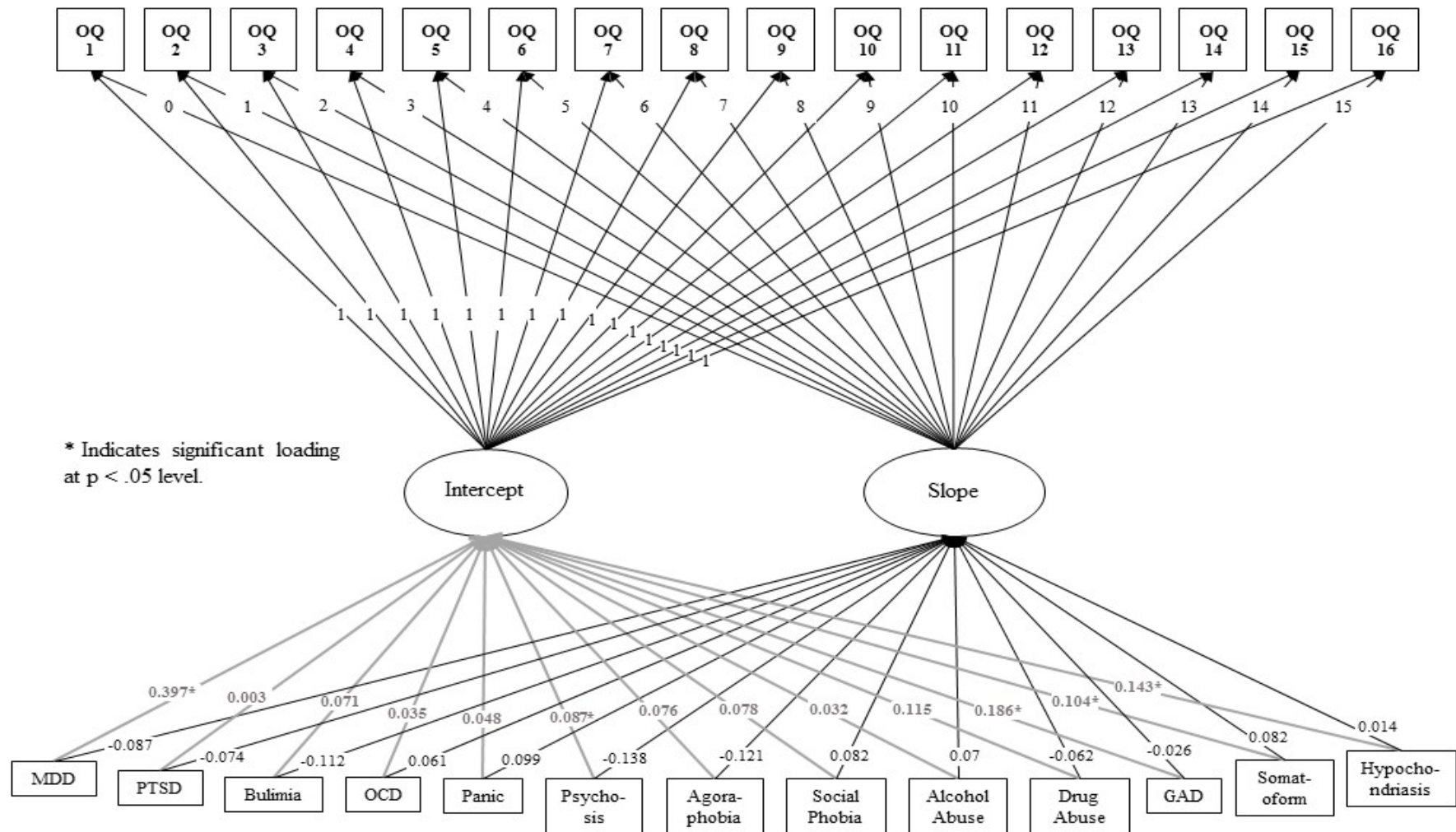


Figure 2. GMM model with PDSQ scores predicting OQ total scores over time.

REFERENCES

- American Psychiatric Association. (1952). *Diagnostic and statistical manual of mental disorders*. Washington, DC: Author.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Bastra, L. & Francis, A. (2012). Diagnostic inflation: Causes and a suggested cure. *Journal of Nervous and Mental Disease*, 200, 474-479. doi: 10.1097/NMD.0b013e318257c4a2
- Beck, A. T. (1967). *Depression: Causes and treatment*. Philadelphia: University of Pennsylvania Press.
- Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. New York, NY: International Universities Press.
- Berkman, L. F. (2000). Social support, social networks, social cohesion and health. *Social Work in Health Care*, 31, 3-14, doi: 10.1300/J010v31n02_02
- Bird S., Klein E., & Loper E. (2009). *Natural language processing with Python*. Sebastopol, California: O'Reilly Media.
- Blashfield, R. K., & Burgess, D. R. (2007). Classification provides an essential basis for organizing mental disorders. In S. O. Lilienfeld & W. T. O'Donohue (Eds.), *The great ideas of clinical science: 17 principles that every mental health professional should understand* (pp. 93-117). New York, NY: Routledge/Taylor & Francis.
- Boyd J.H., Burke J.D. Jr., Gruenberg E., Holzer C.E. 3rd, Rae D.S., George L.K., Karno M., Stoltzman R., McEvoy L., Nestadt G. (1984). Exclusion criteria of DSM-III. A study of co-occurrence of hierarchy-free syndromes. *Archives of General Psychiatry*, 41(10), 983-989. doi: 10.1001/archpsyc.1984.01790210065008.
- Bradford, R. (2008). An empirical study of required dimensionality for large-scale latent semantic indexing applications. *Proceedings of the 17th ACM Conference on Information and Knowledge Management*, 153-162.
- Burgoyne, R. W. (2005). Exploring direction of causation between social support and clinical outcome for HIV-positive adults in the context of highly active antiretroviral therapy. *AIDS Care*, 17, 111-124. doi: 10.1080/09540120412331305179
- Clark, L. A. & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7, 309-319. doi: 10.1037/1040-3590.7.3.309
- Clarkin, J. F. & Kendall, P. C. (1992). Comorbidity and treatment planning; Summary and future directions. *Journal of Consulting and Clinical Psychology*, 60, 904-908. doi: 10.1037/0022-006X.60.6.904

- Clementz, B. A., Sweeney, J. A., Hamm, J. P., Ivleva, E. I., Ethridge, L. E., Pearlson, G. D., ..., & Tamminga, C. A. (2015). Identification of distinct psychosis biotypes using brain-based biomarkers. *American Journal of Psychiatry*, 1173, 373-384. doi: 10.1176/appi.ajp.2015.14091200
- Cloninger, C. R. (2009). Forward. In O'Donohue, W., Fowler, K. A., & Lilienfeld, S. O. (Eds.), *Personality disorders: Toward the DSM-V* (pp. vii-xv). Los Angeles, CA: Sage.
- Cobb, S. (1976). Social support as a moderator of life stress. *Psychosomatic Medicine*, 38, 300-314. doi: 10.1097/00006842-197609000-00003
- Craighead, W. E., Sheets, E. S., Craighead, L. W., & Madsen, J. W. (2011). Recurrence of MDD: A prospective study of personality pathology and cognitive distortions. *Personality Disorders: Theory, Research, and Treatment*, 2(2), 83-97. doi: 10.1037/a0020456
- Cuthbert, B. N. (2014). Translating intermediate phenotypes to psychopathology: The NIMH Research Domain Criteria. *Psychophysiology*, 51(12), 1205-1206. doi: 10.1111/psyp.12342
- Dadds, M. R. & McHough, T. A. (1992). Social support and treatment outcome in behavioral family therapy for child conduct problems. *Journal of Consulting and Clinical Psychology*, 60, 252-259. doi: 10.1037//0022-006X.60.2.252
- Deerwester S., (1988). Improving information retrieval with latent semantic indexing. *Proceedings of the 51st annual meeting of the American Society for Information Science*.
- Dobkin, P. L., Civita, M. D., Paraherakis, A., & Gill, K. (2002). The role of functional social support in treatment retention and outcomes among outpatient adult substance abusers. *Addiction*, 97, 347-356. doi: 10.1046/j.1360-0443.2002.00083.x
- Franklin, J. C., Jamieson, J. P., Glenn, C. R., & Nock, M. K. (2015). How developmental psychopathology theory and research can inform the research domain criteria (RDoC) project. *Journal of Clinical Child & Adolescent Psychology*, 44, 280-290. doi: 10.1080/15374416.2013.873981
- Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice*, 9, 329-343. doi: 10.1093/clipsy.9.3.329
- Helzer, J. E., & Pryzbeck, T. R. (1988). The co-occurrence of alcoholism with other psychiatric disorders in the general population and its impact on treatment. *Journal of Studies on Alcohol*, 49, 219-224. doi: 10.15288/jsa.1988.49.219
- Kaufman, J., Gelernter, J., Hudziak, J. J., Tyrka, A. R., & Coplan, J. D. (2015). The research domain criteria (RDoC) project and studies of risk and resilience in maltreated children. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54, 617-625. doi: 10.1016/j.jaac.2015.06.001

- Kendell, R. & Jablensky, A. (2003). Distinguishing between the validity and utility of psychiatric diagnoses. *American Journal of Psychiatry*, 160, 4-12. doi: 10.1176/appi.ajp.160.1.4
- Kessler, R. C., Avenevoli, S., Costello, E. J., Georgiades, K., Green, J. G., Gruber, M. J., ... & Sampson, N. A. (2012). Prevalence, persistence, and sociodemographic correlates of DSM-IV disorders in the National Comorbidity Survey Replication Adolescent Supplement. *Archives of General Psychiatry*, 69, 372-380. doi: 10.1001/archgenpsychiatry.2011.160.
- Kiss T., Strunk J. (2006). Unsupervised multilingual sentence boundary detection. *Computational Linguist*, 32, 485-525. doi: 10.1162/coli.2006.32.4.485
- Kohlschütter C., Fankhauser P., Nejdl W. (2010). Boilerplate detection using shallow text features. *Proceedings of the Third ACM international conference on Web Search and Data Mining*. New York: Association for Computing Machinery, 441-450. doi: 10.1145/1718487.1718542
- Koons, C. R., Robins, C. J., Tweed, J. L., Lynch, T. R., Gonzalez, A. M., Morse, J. Q., ... & Bastian, L. A. (2001). Efficacy of dialectical behavior therapy in women veterans with borderline personality disorder. *Behavior Therapy*, 32, 371-390. doi: 10.1016/S0005-7894(01)80009-5
- Kraepelin, E. (1899) *PsychiatrieL Ein Lehrbuch* (6th ed.). Leipzig Germany: Barth.
- Lambert, M. J., Gregersen, A. T., & Burlingame, G. M. (2004). The Outcome Questionnaire-45. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment: Volume 3: Instruments for adults* (3rd ed., pp. 191-234). Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.
- Liebllich, S. M., Castle, D. J., & Everall, I. P. (2015). RDoC: We should look before we leap. *Australian and New Zealand Journal of Psychiatry*, 0004867415592956.
- Lilienfeld, S. O., Smith, S. F., & Watts, A. L. (2013). Issues in diagnosis: Conceptual issues and controversies. In E.W. Craighead, D.J. Miklowitz & L. W., Craighead (Eds.), *Psychopathology: History diagnosis, and empirical foundations* (2nd ed., pp. 285-343). Hoboken, NJ: Wiley.
- Loevinger, J. (1957) Objective tests as instruments of psychological theory. *Psychological Reports*, 3, 635-694. doi: 10.2466/pr0.1957.3.3.635
- Long, J. S. and Freese, J. (2006) *Regression Models for Categorical and Limited Dependent Variables Using Stata* (2nd ed.). College Station, Texas: Stata Press.
- Longabaugh, R., Beattie, M., Noel, N., Stout, R., & Malloy, P. (1993). The effect of social investment on treatment outcome. *Journal of Studies on Alcohol*, 54, 465-478. doi: 10.15288/jsa.1993.54.465

- Lynskey, M. T. (1998). The comorbidity of alcohol dependence and affective disorders: treatment implications. *Drug and Alcohol Dependence*, 52, 201-209. doi: 10.1016/S0376-8716(98)00095-7
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7(1), 19-40. doi: 10.1037/1082-989X.7.1.19
- Maj, M. (2005). 'Psychiatric comorbidity': an artefact of current diagnostic systems? *The British Journal of Psychiatry*, 186, 182-184. doi: 10.1192/bjp.186.3.182
- Marcus, D. K., Norris, A. L., & Coccaro, E. F. (2012). The latent structure of attention deficit/hyperactivity disorder in an adult sample. *Journal of Psychiatric Research*, 46, 782-789. doi: 10.1016/j.jpsychires.2012.03.010
- Marcus M. P., Santorini B., & Marcinkiewicz M. A. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguist*, 19(2), 313-330.
- Markon, K. E., Chmielewski, M., & Miller, C. J., (2011). The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychological Bulletin*, 137, 856-879. doi: 10.1037/a0023678
- Mattmann C., & Zitting J. (2011). *Tika in action*. Greenwich, Conn: Manning Publications.
- Maxwell, S. E., & Delaney, H. D. (1993). Bivariate median splits and spurious statistical significance. *Psychological Bulletin*, 113, 181-190. doi: 10.1037/0033-2909.113.1.181
- McCoy, T. H., Castro, V. M., Rosenfield, H. R., Cagan, A., Kohanne, I. S., Perlis, R. H. (2015). A clinical perspective on the relevance of research domain criteria in electronic health records. *American Journal of Psychiatry*, 172, 316-320. doi: 10.1176/appi.ajp.2014.14091177.
- Meehl, P. E., & Golden, R. (1982). Taxometric methods. In P.C. Kendall & J.N. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 127-181). New York, NY: Wiley.
- National Institutes of Mental Health (2013). *Development and definitions of the RDoC domains and constructs*. Retrieved from: <https://www.nimh.nih.gov/research-priorities/rdoc/development-and-definitions-of-the-rdoc-domains-and-constructs.shtml>
- O'hara, M. W., Stuart, S., Gorman, L. L., & Wenzel, A. (2000). Efficacy of interpersonal psychotherapy for postpartum depression. *Archives of General Psychiatry*, 57, 1039-1045. doi: 10.1001/archpsyc.57.11.1039
- Regier, D. A., Narrow, W. E., Clarke, D. E., Kraemer, H. C., Kuramoto, S. J., Kuhl, E. A., & Kupfer, D. J. (2013). *DSM-5 field trials in the United States and Canada, Part II: Test-retest reliability of selected categorical diagnoses*. *American Journal of Psychiatry*, 170, 59-70. doi: 10.1176/appi.ajp.2012.12070999

- Rehurek, R. (2011). Subspace tracking for latent semantic analysis, in ECIR '11. *Proceedings of the 33rd European Conference on Advances in Information Retrieval*, 289-300. doi: 10.1007/978-3-642-20161-5
- Ritshcel, L. A., Gillespie, C. F., Arnarson, E. O., & Craighead, W. E. (2013). Major depressive disorder. In E.W. Craighead, D.J. Miklowitz & L.W. Craighead (Eds.), *Psychopathology: History diagnosis, and empirical foundations* (2nd ed., pp. 285-343). Hoboken, NJ: Wiley.
- Robins, E. & Gauze, S., B. (1970). Establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. *American Journal of Psychiatry*, 126, 983-987. doi: 10.1176/ajp.126.7.983
- Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. *Communications of the ACM*, 18, 613-620. doi: 10.1145/361219.361220
- Schermelleh-Engel, K. & Moosbrugger, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8, 23-74.
- Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death*. San Francisco, CA: Jossey-Bass.
- Singhal, A. (2001). Modern information retrieval: A brief overview. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 24, 35-43.
- Skeem, J. L., Polaschek, D. L., Patrick, C. J., & Lilienfeld, S. O. (2011). Psychopathic personality bridging the gap between scientific evidence and public policy. *Psychological Science in the Public Interest*, 12, 95-162. doi: 10.1177/1529100611426706
- Skinner, B. F. (1953). *Science and human behavior*. New York, NY: Macmillan.
- Sommers C. H., & Satel, S. (2005). *One nation under therapy: How the helping culture is eroding self-reliance*. New York, NY: St. Martin's.
- Steketee, G. (1993). Social support and treatment outcome of obsessive compulsive disorder at 9-month follow-up. *Behavioural and Cognitive Psychotherapy*, 21, 81-95. doi: 10.1017/S014134730001805X
- Vermeersch, D. A., Lambert, M. J., & Burlingame, G. M. (2010). Outcome Questionnaire: Item sensitivity to change. *Journal of Personality Assessment*, 74, 242-261. doi: 10.1207/S15327752JPA7402_6
- Vermeersch, D. A., Whipple, J. L., Lambert, M. J., Hawkins, E. J., Burchfield, C. M., & Okiishi, J. C. (2004). Outcome Questionnaire: Is it sensitive to changes in counseling center clients? *Journal of Counseling Psychology*, 51, 38-49. doi: 10.1037/0022-0167.51.1.38

- Wakefield, J. C. (1998). The *DSM's* theory-neutral nosology is scientifically progressive: Response to Follette and Houts. *Journal of Consulting and Clinical Psychology*, 66, 846-852. doi: 10.1037/0022-006X.66.5.846
- Wakschlag, L. S., Estabrook, R., Petittclerc, A., Henry, D., Burns, J. L., Perlman, S. B., ... & Briggs-Gowan, M. L. (2015). Clinical implications of a dimensional approach: the normal: abnormal spectrum of early irritability. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54(8), 626-634. doi: 10.1016/j.jaac.2015.05.016
- Weston, R. & Gore, P. A. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist*, 34, 719-751. doi: 10.1177/0011000006286345
- Zimmerman, M. (2002). *The Psychiatric Diagnostic Screening Questionnaire*. Los Angeles: Western Psychological Services.
- Zimmerman, M., & Mattia, J. I. (2001). The Psychiatric Diagnostic Screening Questionnaire: Development, reliability and validity. *Comprehensive Psychiatry*, 42, 175-189. doi: 10.1053/comp.2001.23126